

日本国特許庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

Y. INOUE, K. HIRAKAWA  
J. IKEDA

M2047-6



別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office.

出願年月日

Date of Application:

1999年10月 1日

出願番号

Application Number:

平成11年特許願第281716号

出願人

Applicant(s):

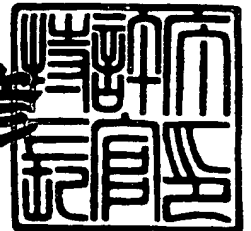
松下電器産業株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2000年 6月23日

特許庁長官  
Commissioner,  
Patent Office

近藤 隆彦



【書類名】 特許願  
【整理番号】 2038610021  
【提出日】 平成11年10月 1日  
【あて先】 特許庁長官殿

【国際特許分類】 H03M 1/22  
H04N 5/92  
H04N 5/93

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式  
会社内

【氏名】 井上 由紀子

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式  
会社内

【氏名】 有村 耕治

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式  
会社内

【氏名】 池田 淳

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100097179

【弁理士】

【氏名又は名称】 平野 一幸

【手数料の表示】

【予納台帳番号】 058698

【納付金額】 21,000円

【提出物件の目録】

【物件名】	明細書	1
【物件名】	図面	1
【物件名】	要約書	1
【プルーフの要否】	要	

【書類名】 明細書

【発明の名称】 圧縮動画像のシーンチェンジ検出装置、圧縮動画像のシーンチェンジ検出方法及びそのプログラムを記録した記録媒体

【特許請求の範囲】

【請求項 1】 入力圧縮動画像の画像構造を判定する画像構造判定部と、前記画像構造判定部の判定結果がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に上下 2 倍分のブロックデータに基づいて特徴量を抽出する特徴量抽出部と、前記特徴量抽出部で抽出されたブロックデータを記録する記憶領域と、抽出されたブロックデータを比較して映像の変化量を求める抽出データ比較部と、前記抽出データ比較部において求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定部を有することを特徴とする圧縮動画像のシーンチェンジ検出装置。

【請求項 2】 入力圧縮動画像の構造画像に関係無く、1 ブロック分のブロックデータに基づいて特徴量を抽出する特徴量抽出部と、前記特徴量抽出部で抽出されたブロックデータを記録する記憶領域と、特徴量が抽出された画像がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に 2 倍のブロックデータを用いて特徴量を比較し変化量を求める抽出データ比較部と、前記抽出データ比較部において求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定部を有することを特徴とする圧縮動画像のシーンチェンジ検出装置。

【請求項 3】 画像がフレーム構造画像である場合、フィールド DCT 符号化されたブロック数をカウントするフィールド DCT 符号化ブロック数カウント部と、このブロック数と閾値とを比較して、フィールド間に存在するシーンチェンジを判定するシーンチェンジ判定部とを有することを特徴とする圧縮動画像のシーンチェンジ検出装置。

【請求項 4】 シーンチェンジを判定するシーンチェンジ判定部と、前記シーンチェンジ判定部が検出したシーンチェンジのうち、指定された特定の間隔の開始点と終了点に存在するシーンチェンジを検索するシーンチェンジ間隔検索部とを有することを特徴とする圧縮動画像のシーンチェンジ検出装置。

【請求項 5】前記シーンチェンジ判定部が、シーンチェンジの判定基準として用いる閾値には、画像の最大変化量を基準として定められる閾値が含まれていることを特徴とする請求項 1 から 4 記載の圧縮動画像のシーンチェンジ検出装置。

【請求項 6】入力圧縮動画像の画像構造を判定する画像構造判定ステップと、前記画像構造判定ステップの判定結果がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に上下 2 倍分のブロックデータに基づいて特徴量を抽出する特徴量抽出ステップと、前記特徴量抽出ステップで抽出されたブロックデータを記録する記憶領域と、抽出されたブロックデータを比較して映像の変化量を求める抽出データ比較ステップと、前記抽出データ比較ステップにおいて求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定ステップを有することを特徴とする圧縮動画像のシーンチェンジ検出方法。

【請求項 7】入力圧縮動画像の構造画像に関係無く、1 ブロック分のブロックデータに基づいて特徴量を抽出する特徴量抽出ステップと、前記特徴量抽出ステップで抽出されたブロックデータを記録する記憶領域と、特徴量が抽出された画像がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に 2 倍のブロックデータを用いて特徴量を比較し変化量を求める抽出データ比較ステップと、前記抽出データ比較ステップにおいて求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定ステップを有することを特徴とする圧縮動画像のシーンチェンジ検出方法。

【請求項 8】画像がフレーム構造画像である場合、フィールド DCT 符号化されたブロック数をカウントするフィールド DCT 符号化ブロック数カウントステップと、このブロック数と閾値とを比較して、フィールド間に存在するシーンチェンジを判定するシーンチェンジ判定ステップとを有することを特徴とする圧縮動画像のシーンチェンジ検出方法。

【請求項 9】シーンチェンジを判定するシーンチェンジ判定ステップと、前記シーンチェンジ判定ステップにおいて検出したシーンチェンジのうち、指定された特定の間隔の開始点と終了点に存在するシーンチェンジを検索するシーンチェンジ間隔検索ステップとを有することを特徴とする圧縮動画像のシーンチェンジ検出方法。

【請求項 1 0】前記シーンチェンジ判定ステップにおいて、シーンチェンジの判定基準として用いる閾値には、画像の最大変化量を基準として定められる閾値が含まれていることを特徴とする請求項 6 から 9 記載の圧縮動画画像のシーンチェンジ検出方法。

【請求項 1 1】入力圧縮動画画像の画像構造を判定する画像構造判定ステップと、前記画像構造判定ステップの判定結果がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に上下 2 倍分のブロックデータに基づいて特徴量を抽出する特徴量抽出ステップと、前記特徴量抽出ステップで抽出されたブロックデータを記録する記憶領域と、抽出されたブロックデータを比較して映像の変化量を求める抽出データ比較ステップと、前記抽出データ比較ステップにおいて求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定ステップを有することを特徴とする圧縮動画画像のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体。

【請求項 1 2】入力圧縮動画画像の構造画像に関係無く、1 ブロック分のブロックデータに基づいて特徴量を抽出する特徴量抽出ステップと、前記特徴量抽出ステップで抽出されたブロックデータを記録する記憶領域と、特徴量が抽出された画像がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に 2 倍のブロックデータを用いて特徴量を比較し変化量を求める抽出データ比較ステップと、前記抽出データ比較ステップにおいて求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定ステップを有することを特徴とする圧縮動画画像のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体。

【請求項 1 3】画像がフレーム構造画像である場合、フィールド D C T 符号化されたブロック数をカウントするフィールド D C T 符号化ブロック数カウントステップと、このブロック数と閾値とを比較して、フィールド間に存在するシーンチェンジを判定するシーンチェンジ判定ステップとを有することを特徴とする圧縮動画画像のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体。

【請求項 1 4】シーンチェンジを判定するシーンチェンジ判定ステップと、前

記シーンチェンジ判定ステップにおいて検出したシーンチェンジのうち、指定された特定の区隔の開始点と終了点に存在するシーンチェンジを検索するシーンチェンジ区隔検索ステップとを有することを特徴とする圧縮動画像のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体。

【請求項 1 5】前記シーンチェンジ判定ステップにおいて、シーンチェンジの判定基準として用いる閾値には、画像の最大変化量を基準として定められる閾値が含まれていることを特徴とする請求項 1 1 から 1 4 記載の圧縮動画像のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、圧縮動画像から、シーンの変わり目を検出するシーンチェンジ検出装置及びその関連技術に関するものである。

【0 0 0 2】

【従来の技術】

近年、デジタルビデオカメラの普及や、デジタル放送の出現などにより、MPEGやDVなど、圧縮動画像を扱う機会が増加している。また、過去の大量なアナログ映像を、デジタルの圧縮動画像として保存する場合もある。そして、このような圧縮動画像を復号せずに、符号化されたまま、編集する技術が実用化されつつある。

【0 0 0 3】

このような編集においては、圧縮動画像、すなわちビットストリームのなかから、シーンチェンジ（映像又は場面の変わり目の位置）を、高速に自動検出する技術が是非とも必要となる。なぜなら、検出されたシーンの先頭の位置情報やシーンチェンジ技術によって切り分けられたシーンの代表画像は、映像内容のインデックスとして有用であり、内容の検索や編集の重要な手助けとなるからである。

【0 0 0 4】

次に、圧縮動画像のフォーマットとして広く用いられている、MPEG 2 の符

号化について説明する。MPEG 2 符号化は、MPEG 1, H. 261 と同様に動きベクトルと DCT (離散コサイン変換: Discrete Cosine Transform) を用いた方式である。データは、フレーム内で、輝度 (Y)、色差 (Cb, Cr) に分離され、16 \* 16 画素のマクロブロック単位で、符号化される。

## 【0005】

そして、各マクロブロックを符号化するにあたっては、参照画像から動き予測を行う動き補償予測が選択されるか、または、符号化を行うデータのみで符号化を行うイントラ符号化が選択される。

## 【0006】

このうち、動き補償予測は、フレーム間の時間的な相関が高い場合に符号化率が高くなる方式であり、符号化を行うマクロブロックと、参照画像から動き予測によって得られるマクロブロックのデータとの、差分から予測誤差信号を得て時間空間的に情報の圧縮を行うものである。なお、動き補償予測では、この予測誤差信号が、8 \* 8 画素のブロック単位で、DCT により空間周波数領域に変換される。

## 【0007】

一方、イントラ符号化は、符号化されるブロックのデータそのものを、8 \* 8 画素のブロックに分け、これらのブロック単位で、単純に DCT 符号化する方式である。

## 【0008】

以下、符号化の単位について説明する。MPEG 2 では、インタレース画像も対象としており、画面の符号化の単位として、フレーム構造とフィールド構造とがある。

## 【0009】

フレーム構造では、奇数フィールドと偶数フィールドの 2 フィールドをインタレースに配置したフレームに対して符号化を行う。一方、フィールド構造では、奇数フィールドもしくは偶数フィールドの、1 フィールドに対して符号化を行う。



## 【0 0 1 0】

さて、本明細書において、フレーム構造で符号化された画像を「フレーム構造画像」といい、フィールド構造で符号化された画像を「フィールド構造画像」という。

## 【0 0 1 1】

次に、動き補償について説明する。MPEG 2では、上述のように、フレーム構造とフィールド構造とがある。そして、フレーム構造画像の動き補償予測については、フレーム予測、フィールド予測及びデュアルプライム (Dual-Prime) 予測がある。また、フィールド構造画像の動き補償予測としては、フィールド予測、16 \* 8MC予測及びデュアルプライム予測がある。さらに、フレーム予測以外の予測においては、参照するフィールドが奇数フィールドであるか偶数フィールドであるかを選択することができる。

## 【0 0 1 2】

次に、図15を参照しながら、符号化の方法について説明する。フレーム構造画像では、符号化の方法として、フレームDCTとフィールドDCTという2種類のDCTを用いることができる。

## 【0 0 1 3】

このうち、フレームDCTは、マクロブロックの輝度信号を、図15 (a) に示すように、4個のブロックに分解する際に、各ブロックがフレームで構成されるように分解し、これにDCTを施すものである。

## 【0 0 1 4】

一方、フィールドDCTは、図15 (b) に示すように、マクロブロックの輝度信号を4個のブロックに分解する際に、各ブロックがフィールドで構成されるように分解し、これにDCTを施すものである。

## 【0 0 1 5】

さて、符号化の際には、この2種類のDCTのどちらを用いてもよいが、一般的に、奇数フィールドと偶数フィールドの画像データの差が大きい場合にフィールドDCTを用いると符号化の効率がよくなることが知られている。特に、二つのシーンが、一つのフィールドに混在しているような場合は、フィールドDCT

を行うと圧縮率が高くなる。

【0 0 1 6】

しかし、フィールドDCTの場合、フレーム構造を2つのフィールドに分解する必要があるため、フレームDCTに対して処理速度は遅くなる。つまり、このような性質に合わせて、2種類のDCTをうまく用いることにより、フレーム構造画像（インタレース画像）の符号化効率を向上させることができる。なお、4 : 2 : 0フォーマットにおける色差信号については、常にフレームDCTを用いる。また、フィールド構造画像では、マクロブロックが1フィールドの信号のみで構成されるため、常にフィールドDCTを行うことになる。

【0 0 1 7】

以上の説明をふまえて、以下従来のシーンチェンジ技術について説明する。従来、シーンチェンジ検出には、（1）画像の色のヒストグラム、（2）圧縮動画のデータサイズ、（3）2フレームの画像の同一位置におけるブロックデータ差分、などの特徴量を用いている。

（1）画像の色のヒストグラムを用いるには、1フレームの画像に用いられている色を、1フレーム分あるいは1フレームを分割した領域において、ヒストグラムにし、このヒストグラムをそのフレームの特徴量とし、その前後のフレーム画像における特徴量と比較して類似度を求める（例えば、特開平7-59108号公報参照）。

（2）圧縮動画のデータサイズを用いるには、シーンチェンジ部分では、圧縮率が悪いという性質を利用して、隣り合うフレームのデータのサイズを比較し、その差分が所定の閾値よりも大きくなった時にシーンチェンジと判定する（例えば、特開平7-121555号公報参照）。

【0 0 1 8】

しかし、これら（1）～（2）の手法では、フレーム単位でしかシーンチェンジを検出できないから、シーンチェンジが、1フレーム中の奇数フィールドと偶数フィールドの間（2つのフィールド間）に存在すると、このシーンチェンジを正確に検出できない。

【0 0 1 9】

この問題を解決するために、特開平 9 - 3 2 2 1 2 0 号公報では、フィールド予測方式を用いて符号化された映像データから復号処理を行うこと無しにシーンチェンジを検出する方法を提案している。このものでは、予測が行われるフレームにおいて、参照フレームの奇数フィールドもしくは偶数フィールドのどちらを選択して予測を行うかという、参照フィールド選択信号に基づいて、フィールド間の類似度を複数算出し、その結果を用いてシーンチェンジ検出を行っている。

## 【 0 0 2 0 】

## 【発明が解決しようとする課題】

しかしながら、この手法は、フィールド予測方式に依存しているから、フィールド予測方式（フレーム間予測方式）を用いない映像や、フィールド予測方式の映像とそれ以外の予測方式の映像が混在する映像に対しては、適用できない。

## 【 0 0 2 1 】

また、（3）データとして同じ位置の D C T における D C 成分のみの差分を用いると、位置の対応がとれない場合がある。これは、フレーム構造画像では、符号化の方法としてフレーム D C T とフィールド D C T という 2 種類の D C T を用いることができるため、D C T を解かずにブロックデータの比較を行うと、比較する片方のデータがフレーム D C T で符号化され、もう一方のデータがフィールド D C T で符号化されていた場合に、画像における  $8 * 8$  画素のデータと、 $8 * 16$  画素データ中の、奇数もしくは偶数フィールドのみの  $8 * 8$  画素のデータとを、比較することになるからである。

## 【 0 0 2 2 】

これを改善するには、フレーム構造画像 1 フレーム分と、フィールド構造画像 1 フレーム相当分（奇数フィールド画像と偶数フィールド画像）とを比較しなければならない。しかし、これでは、2 枚のフィールド構造画像のデータが揃った時点でしか比較ができず、処理が煩雑で処理速度も遅くなるという、問題点がある。

## 【 0 0 2 3 】

そこで本発明は、フィールド／フレームの各構造が混在していても、シーンチェンジを検出可能な技術を提供することを、第 1 の目的とする。

【 0 0 2 4 】

また、本発明は、フレーム構造のフィールド間に存在するシーンチェンジをも検出できる技術を提供することを、第 2 の目的とする。

【 0 0 2 5 】

さらに、本発明は、予め開始点と終了点との間隔が分かっている場合に、迅速に目的のシーンチェンジを検出できる技術を提供することを、第 3 の目的とする。

【 0 0 2 6 】

【課題を解決するための手段】

本発明は、第 1 の目的のため、請求項 1、2、6、7、11、12 において、圧縮動画像がフレーム構造画像の場合には、フィールド構造画像に対応する縦方向に 2 倍のデータを抽出する。

【 0 0 2 7 】

本発明は、第 2 の目的のため、請求項 3、8、13 において、フィールド DCT 符号化ブロック数カウント部を備える。

【 0 0 2 8 】

本発明は、第 3 の目的のため、請求項 4、9、14 において、複数の検出されたシーンチェンジの中から、指定された間隔に一致するシーンチェンジの組を検出し、そのシーンチェンジを結果として出力する。

【 0 0 2 9 】

【発明の実施の形態】

本発明におけるシーンチェンジ検出は、全て、予測方式に依存しない。したがって、予測方式が何であれ、あるいは、予測方式の有無にかかわらず、シーンチェンジを検出できる。なお以下、本発明の全ての実施の形態において、入力圧縮動画像は、MPEG 2 によることとして説明を行うが、DCT を用い、かつ、フィールド／フレームの各構造を混在させ得る圧縮方式であれば、同様の効果が得られるから、このような圧縮方式による技術であれば、本発明の対象に包含される。

【 0 0 3 0 】

## (第 1 の実施の形態)

以下、図面を参照しながら、本発明の実施の形態を説明する。図 1 は、本発明の第 1 の実施の形態におけるシーンチェンジ検出装置のブロック図である。

## 【0031】

図 1 に示すように、このシーンチェンジ検出装置には、MPEG2 に従って符号化されたビットストリームが、入力圧縮動画像として入力される。また、このシーンチェンジ検出装置からは、ビットストリームにおいて検出されたシーンチェンジの位置を示す情報が、検出結果として出力される。ここで、従来の技術の項で述べたように、入力圧縮動画像の符号化方式、符号化の単位（フレーム／フィールド構造）、DCT などは、種々のものが考えられ、また、1 種類だけでなく、複数種のものが、時間軸上で混在していてもよい。

## 【0032】

就中、フィールド構造画像とフレーム構造画像とが、時間軸上で交互に存在するような場合でも差し支えない。以上の点は、後述する他の実施の形態についても、そのまま適合する。

## 【0033】

図 1 に示すように、入力圧縮動画像は、まず、画像構造判定部 1 に入力される。画像構造判定部 1 は、現在入力している画像が、フィールド構造画像、フレーム構造画像のいずれであるかを、ビットストリームの特定領域の情報を参照して、判定する。そして、この判定結果と、ビットストリームの内容が、次段の特徴量抽出部 2 に出力される。

## 【0034】

次に、特徴量、変化量及び特徴量抽出部 2 の動作を、図 5 ～図 6 を例にとって、説明する。図 5 では、図の左側が時間軸上で古い画像を示し、右側が新しい画像を示し、 $t$  枚目の画像を、画像  $t$  のように記述する。また、この例では、画像  $t-2 \sim t+1$ 、 $t+4 \sim t+5$  がフィールド構造画像であり、画像  $t+2 \sim t+3$  はフレーム構造画像である。即ち、画像  $t+1 \sim t+2$ 、 $t+3 \sim t+4$  では、構造の変化がある。

## 【0035】

そして、図 6 (a) (トップフィールド)、図 6 (b) (ボトムフィールド) のような、フィールド構造画像では、1 ブロック分のブロックデータ  $D_t$  を用い、ブロックデータ  $D_t$  をそのまま特徴量  $d_t$  として、使用する ( $d_t = D_t$ )。

【0036】

また、図 6 (c) のような、フレーム構造画像では、上下 2 ブロック分のブロックデータ (ブロックデータ  $D_{tu}$  とブロックデータ  $D_{tb}$ ) を用い、その平均値を特徴量  $d_t$  として使用する ( $d_t = (D_{tu} + D_{tb}) / 2$ )。

【0037】

ここで、ブロックデータ  $D_t$ 、 $D_{tu}$ 、 $D_{tb}$  は、同種のデータでなければならないが、ブロック内の輝度平均値やその他ブロック内の画像を代表する種々のデータを用いることができる。

【0038】

また、図 6 (a)、(b)、(c) に示しているように、フレーム構造画像は、フィールド構造画像に対して、上下 2 倍の高さを持つので、フィールド構造画像とフレーム構造画像とを比較するには、フィールド構造画像において、ブロック座標  $(x, y)$  のブロックデータ  $D_t$  と、ブロック座標  $(x, 2 * y)$  のブロックデータ  $D_{tu}$  及びブロック座標  $(x, 2 * y + 1)$  のブロックデータ  $D_{tb}$  との組とを、対比して用いる。

【0039】

次に、時間軸上で異なる位置にある、画像  $t$  と画像  $s$  (典型的には、 $s = t + 1$ ) とにおける、ブロック座標  $(x, y)$  の特徴量を、それぞれ  $d_t(x, y)$ 、 $d_s(x, y)$  とするとき、画像  $t$  と画像  $s$  とにおける、第 1 の変化量  $R(t, s)$  を次のように定義する。この第 1 の変化量  $R(t, s)$  は、画像  $t$  と画像  $s$  との単純な差分に近い。

【0040】

【数 1】

$$R(t, s) = \sum_{x, y} | d_t(x, y) - d_s(x, y) |$$

【0 0 4 1】

また、同様に、第2の変化量 $Q(t, s)$ を次のように定義する。第2の変化量 $Q(t, s)$ は、画像 $t$ と画像 $s$ の微分値に近い。

【0 0 4 2】

【数2】

$$Q(t, s) = 2R(t, s) - R(t-1, t-1) - R(t+1, t+1)$$

【0 0 4 3】

ここで、第1の変化量 $R(t, s)$ と、これに対して、予め経験的に設定される第1の閾値とを、比較することにより、画像に大きな変化が現れた位置を取り出すことができ、この位置をシーンチェンジと検出できる。但し、画像によっては、動きの激しいものなど、シーンは同じではあるが、連続的な変化が起き、第1の変化量 $R(t, s)$ のみでは、過剰にシーンチェンジを検出するおそれがある。このような場合、第2の変化量 $Q(t, s)$ と、これに対して、予め設定される第2の閾値との、比較を併用することにより、不連続で、かつ、大きな変化がある位置のみを、シーンチェンジと検出することができ、検出されるシーンチェンジの信頼性を向上できる。

【0 0 4 4】

なお、動画像によっては、第1の変化量 $R(t, s)$ 、第2の変化量 $Q(t, s)$ の一方のみで十分な場合もあり得るし、第1の変化量 $R(t, s)$ 、第2の変化量 $Q(t, s)$ に適当な重みつけをして、両方検討するようにしても差し支えない。勿論、以上の2つの式は、単なる例示に過ぎず、これらと等価な式はもとより、シーンチェンジを判定できるものであれば、他の式で代用しても良い。

【0 0 4 5】

これらの第1の変化量 $R(t, s)$ 、第2の変化量 $Q(t, s)$ を求めるには、図3に示すように、入力画像1枚分の抽出データの全てを得た後に、比較対象画像との比較をし変化量を算出しても良いが、図4に示すように、1ブロック（もしくは複数ブロック）のデータを抽出した時に、同時に、比較対象画像の同じ

位置に対応するデータとの比較を行う方が望ましい。なぜなら、図4のようにした方が、処理速度は上がるからである。

【0046】

このようにすると、時間軸上において、フィールド構造画像とフレーム構造画像の混合する一連の映像から、一律にシーンチェンジを検出できる。

【0047】

さて、図1に示すように、特徴量抽出部2が抽出した特徴量 $d_t$ と、特徴量 $d_t$ が抽出されたフレーム番号などの画像情報とが、第1のデータメモリ6に対応付けて格納される。この第1のデータメモリ6は、記憶領域に相当するものであり、ハードディスク装置などの他の記憶手段で代用しても良い。

【0048】

特徴量抽出部2の次段にある抽出データ比較部3は、特徴量抽出部2が新たに特徴量 $d_t$ を抽出し、第1のデータ用メモリ6に、この新たな特徴量 $d_t$ を記録すると、この新たな特徴量 $d_t$ と、それ以前（典型的にはその直前回）に第1のデータ用メモリ6に記録された特徴量とを参照し、これらの特徴量間の変化量を求める。この変化量は、第1の変化量 $R(t, s)$ 、第2の変化量 $Q(t, s)$ の一方のみでも良いが、望ましくは、両方求める。

【0049】

そして、抽出データ比較部3は、求めた変化量を比較結果情報として、第2のデータ用メモリ7に記録すると共に、この変化量が生じたフレーム番号等の画像情報を、比較結果情報に対応付けて第2のデータ用メモリ7に記録する。

【0050】

また、シーンチェンジ判定用データ入力部5は、シーンチェンジ判定時に使用する閾値を保持し、この閾値をシーンチェンジ判定部4へ出力する。この閾値としては、第1の変化量 $R(t, s)$ 用の第1の閾値と、第2の変化量 $Q(t, s)$ 用の第2の閾値とがある。

【0051】

シーンチェンジ判定部4は、抽出データ比較部3が比較を行うと、第2のデータ用メモリ7に記録された比較結果情報を参照し、これがシーンチェンジ判定用



データ入力部 5 から入力する閾値を越えていれば、この位置でシーンチェンジが発生したと判定し、この比較結果情報に係る、フレーム番号等の画像情報（つまり、ビットストリーム内の位置）を検出結果として出力する。また、そうでなければ、シーンチェンジ判定部 4 は、検出結果を出力しないようにしても良いし、あるいは、検出しなかった旨の検出結果を出力するようにしても良い。

## 【 0 0 5 2 】

図 2 は、本発明の第 1 の実施の形態におけるシーンチェンジ検出装置のフローチャートである。次に、図 2 を用いて、本形態のシーンチェンジ検出装置の動作を説明する。まず、入力圧縮動画像が、画像構造判定部 1 に至ると、この判定部 1 は、現在の画像が、フレーム構造画像／フィールド構造画像のいずれであるかを判定する（ステップ 1）。フレーム構造画像であれば、特徴量抽出部 2 は、画像の縦方向上下 2 ブロック分のデータを用いた値を特徴量として抽出し、第 1 のデータ用メモリ 6 に記録する（ステップ 2）。フィールド構造画像であれば、特徴量抽出部 2 は、1 ブロック分のデータを用いた値を特徴量として抽出し、第 1 のデータ用メモリ 6 に記録する（ステップ 3）。つまり、フレーム構造画像ではフィールド構造画像の縦方向 2 倍のデータを用いる。

## 【 0 0 5 3 】

そして、ステップ 4 にて、抽出データ比較部 3 が、今回の特徴量と、それ以前の特徴量とを比較して、比較結果情報を第 2 のデータ用メモリ 7 に記録する。次に、ステップ 5 にて、シーンチェンジ判定部 4 は、この比較結果情報をシーンチェンジ判定用データ入力部 5 から入力する、閾値と比較し、シーンチェンジと判定できれば、シーンチェンジが発生した位置を検出結果として出力する。

## 【 0 0 5 4 】

なお、本形態では、特徴量  $d_t$  として、フィールド構造画像では、1 ブロック分のブロックデータ  $D_t$  を用い、フレーム構造画像では、縦方向に 2 ブロック分のブロックデータ  $D_{tu}$ 、 $D_{tb}$  の平均値を用いたが、フィールド構造画像とフレーム構造画像のデータの比較レベルが同等になれば、他の手法によっても良い。例えば、フィールド構造画像の 1 ブロック分のブロックデータを 2 倍にしたもの（ $d_t = 2 * D_t$ ）と、フレーム構造画像の縦方向上下 2 ブロックのブロック

データの和 ( $d_t = D_{tu} + D_{tb}$ ) とを、特徴量  $d_t$  として、これらと比較してもよい。

## 【0055】

また、本形態では、フィールド構造画像において抽出されるデータを1ブロックずつの単位とし、フレーム構造画像において抽出されるデータを2ブロックずつの単位としたが、フレーム構造画像から抽出されるデータが、フィールド構造画像において抽出されるデータと対応する位置に存在する縦方向に2倍のデータであれば、同様にシーンチェンジを検出できる。つまり、フィールド構造画像Nブロックのデータを特徴量として抽出した場合、フレーム構造画像からは対応する位置の縦方向に2倍の、 $2 * N$ ブロックからのデータを、特徴量として抽出するようにすればよい。

## 【0056】

これにより、フレーム構造画像とフィールド構造画像の混合する圧縮動画像からも、一律にシーンチェンジを検出できる。また、フレーム構造画像内でフレームDCTを用いられようとフィールドDCTを用いられようと、特別な配慮をすることなくシーンチェンジを検出できる。また、この手法は、予測方式に依存しないので、予測方式が何であるかは不問であるし、予測方式の有無にかかわらず、所望のシーンチェンジを検出できる。

## 【0057】

(第2の実施の形態)

図7は、本発明の第2の実施の形態におけるシーンチェンジ検出装置のブロック図、図8は、本発明の第2の実施の形態におけるシーンチェンジ検出装置のフローチャートである。

## 【0058】

図7では、第1の実施の形態に係る図1に対し、特徴量抽出部2と画像構造判定部1の位置を入れ換えてある。また、本形態の特徴量抽出部2は、入力された圧縮動画像がフレーム構造画像であってもフィールド構造画像であっても、画面全体について、1ブロックずつのブロックデータから、特徴量を抽出する点異なる(ステップ10)。

## 【0059】

そして、図8に示すように、抽出データ比較部3の前段にある、画像構造判定部1は、今回入力した画像の構造を調べ、フレーム構造画像の場合、抽出データ比較部3は、画像の縦方向に上下2ブロック分のデータを用いてデータを比較する（ステップ12）。一方、フィールド構造画像ならば、抽出データ比較部3は、1ブロック分のデータを用いてデータを比較する（ステップ13）。ここでの比較は、第1の実施の形態と同様である。

## 【0060】

さて、図5の例でいえば、画像 $t$ と画像 $t+1$ とを比較するとき、両画像ともフィールド構造画像であるので、画面上の同じ位置にあるブロックのデータが比較される。

## 【0061】

画像 $t+1$ と画像 $t+2$ の比較であれば、画像 $t+2$ はフレーム構造画像であるので、画像 $t+1$ におけるブロック座標 $(x, y)$ から得られるデータと、画像 $t+2$ のデータは、ブロック座標 $(x, 2 * y)$ とブロック座標 $(x, 2 * y + 1)$ から得られるデータとが、比較される。

## 【0062】

そして、例えば、画像 $t+1$ におけるブロック座標 $(x, y)$ から得られるデータをA、ブロック座標 $(x, 2 * y)$ とブロック座標 $(x, 2 * y + 1)$ から得られるデータをそれぞれB、Cとすると、データAとデータB、Cの平均値との差分の絶対値を、変化量とする。この変化量を画像全体において求めることで、2枚の画像の変化量が得られる。

## 【0063】

なお、ここでは、フレーム構造画像のデータとして縦方向に2ブロック分の平均値を用いたが、これは単なる和でも良く、その場合、対応するフィールド構造画像のデータは、2倍にするなど、データの正規化を行えば良い。具体的には、データAの2倍の値とデータB、Cの和との差分の絶対値を変化量とすると良い。

## 【0064】

その他の点は、第 1 の実施の形態と同様である。

【0065】

(第 3 の実施の形態)

図 9 は、本発明の第 3 の実施の形態におけるシーンチェンジ検出装置のブロック図、図 10 は、本発明の第 3 の実施の形態におけるシーンチェンジ検出装置のフローチャートである。

【0066】

本形態では、図 1 と図 9 とを比較すれば明らかなように、フィールド DCT 符号化ブロック数カウント部 8 と、第 3 のデータ用メモリ 9 を追加している。このフィールド DCT 符号化ブロック数カウント部 8 は、画像構造判定部 1 による判定結果が、フレーム構造画像であった場合、その符号化がフレーム DCT を用いているのかフィールド DCT を用いているのかを判定し、1 フィールド中のフィールド DCT 符号化が行われているマクロブロック数（ブロック数でも良い）をカウントする。そして、フィールド DCT 符号化ブロック数カウント部 8 がカウントしたフィールド DCT 符号化ブロック数は、第 3 のデータ用メモリ 9 に記録される。

【0067】

したがって、図 10 に示すように、まず、入力圧縮動画像が、画像構造判定部 1 に至ると、この判定部 1 は、現在の画像が、フレーム構造画像／フィールド構造画像のいずれであるかを判定する（ステップ 20）。フレーム構造画像であれば、特徴量抽出部 2 は、画像の縦方向上下 2 ブロック分のブロックデータを用いた値を特徴量として抽出し、第 1 のデータ用メモリ 6 に記録し（ステップ 21）、フィールド DCT 符号化ブロック数カウント部 8 がフィールド DCT 符号化ブロック数をカウントし、このブロック数が第 3 のデータ用メモリ 9 に記録される（ステップ 23）。なお、ステップ 21、23 の順序は入れ換えても差し支えない。

【0068】

一方、フィールド構造画像であれば、特徴量抽出部 2 は、1 ブロック分のブロックデータを用いた値を特徴量として抽出し、第 1 のデータ用メモリ 6 に記録す

る（ステップ 2 2）。つまり、フレーム構造画像ではフィールド構造画像の縦方向 2 倍のデータを用いる。

## 【 0 0 6 9 】

そして、ステップ 2 4 にて、抽出データ比較部 3 が、今回の特徴量と、それ以前の特徴量とを比較して、比較結果情報を第 2 のデータ用メモリ 7 に記録する。次に、ステップ 2 5 にて、シーンチェンジ判定部 4 は、この比較結果情報をシーンチェンジ判定用データ入力部 5 から入力する、閾値と比較し、シーンチェンジと判定できるかどうか検討する。さらに、ステップ 2 6 にて、シーンチェンジ判定部 4 は、第 3 のデータ用メモリ 9 に記録した、ブロック数と閾値とを比較して、シーンチェンジと判定できるかどうか検討する。そして、シーンチェンジ判定部 4 は、ステップ 2 5 又はステップ 2 6 のいずれかで、シーンチェンジと判定したら、発生した位置を検出結果として出力する。

## 【 0 0 7 0 】

さて、図 1 1 のように、二つのシーンがフィールドで混ざっているような場合には、入力画像と直前直後の画像との変化量が小さくなってしまう場合があり、検出漏れを起こす原因となっていた。しかし、従来の技術の項で述べたように、このような入力画像においてフィールド D C T を行うと圧縮率が高くなる。

## 【 0 0 7 1 】

そのため、フレーム内でフィールド D C T が多く用いられている場合には、フレーム内の奇数フィールドと偶数フィールドの相関が低いと見做すことができる。このため、特徴量抽出部 2 は、フィールド D C T 符号化が行われている数をカウントし、これをフレーム内の第 3 の変化量として、比較検討対象に追加する。

## 【 0 0 7 2 】

抽出データ比較部 3 は、第 1，第 2 の実施の形態と同様であるが、シーンチェンジ判定部 4 は、第 1，第 2 の実施の形態における判定に加えて、第 3 の変化量と、この第 3 の変化量のために予め設定された、第 3 の閾値とを比較した場合に、第 3 の変化量が第 3 の閾値よりも大きい時には、フレームのフィールド間にシーンチェンジがあると判断する。

## 【 0 0 7 3 】

以上、説明したように、フィールド DCT が用いられた数をカウントすることにより、従来検出が非常に困難であった、フレーム構造画像の 2 つのフィールド間に存在するシーンチェンジ（図 11 に例示している）を検出できる。

## 【0074】

さらに、本形態では、シーンチェンジ判定用データ入力部 5 が、シーンチェンジ判定部 4 に出力する閾値について、次の工夫がなされている。即ち、画像の最大変化量を基準（100%）として、閾値は、この基準の所定パーセントと定める。

## 【0075】

例えば、比較画像 A、B において、比較する 1 ブロックのデータの取り得る最小値が 0 であり最大値が 255 であるとする、1 ブロックの最大変化量は 255 である。そして、画像の比較に用いたブロック数が 1320 であるとする、画像全体の最大変化量は、 $255 * 1320 = 336600$  となる。この変化量を基準（100%）とする。また、ここでの閾値は、例えば 3%（10098）～10%（33660）程度が好適である。

## 【0076】

勿論、使用するデータの数やデータの最大変化量が変われば、それに伴って閾値は変化するが、閾値と基準の比率は一定とする。

## 【0077】

これにより、画像サイズ（縦×横）が変わったり、判定に使用されるデータの種類が変わったりしても、検出のばらつきを抑制して、ほぼ一様な検出結果を得ることができる。

## 【0078】

（第 4 の実施の形態）

図 12 は、本発明の第 4 の実施の形態におけるシーンチェンジ検出装置のブロック図、図 13 は、本発明の第 4 の実施の形態におけるシーンチェンジ検出装置のフローチャートである。

## 【0079】

本形態では、図 1 と図 12 とを比較すれば明らかなように、シーンチェンジ判

定部 4 が検出結果をダイレクトに出力するのではなく、シーンチェンジ判定部 4 がシーンチェンジと判定したシーンチェンジ位置情報を、一旦、第 4 のデータ用メモリ 1 1 に格納するようにしている。

【0 0 8 0】

また、シーンチェンジ判定用データ入力部 5 には、基準となるシーンチェンジからターゲットシーンまでの時間軸上の間隔が設定される。そして、シーンチェンジ間隔検索部 1 0 を追加している。このシーンチェンジ間隔検索部 1 0 は、第 4 のデータ用メモリ 1 1 に記録されたシーンチェンジ位置情報同士の時間軸上の間隔を求め、求めた間隔と、シーンチェンジ判定用データ入力部 5 から与えられる間隔とを、比較する。

【0 0 8 1】

したがって、図 1 3 に示すように、先の実施の形態と同様に、シーンチェンジ判定部 4 は、シーンチェンジを探す（ステップ 3 0）。そして、シーンチェンジ判定部 4 が、シーンチェンジを見つけると、このシーンチェンジ位置情報を第 4 のデータ用メモリ 1 1 に格納する（ステップ 3 1）。

【0 0 8 2】

そして、シーンチェンジ間隔検索部 1 0 は、第 4 のデータ用メモリ 1 1 をアクセスして、シーンチェンジ間隔の間隔を調べ（ステップ 3 2）、シーンチェンジ判定用データ入力部 5 から与えられた間隔と一致するシーンチェンジの組が見つかり、見つかったシーンチェンジの先頭と末尾とからなる、組の位置情報を、検出結果として出力する（ステップ 3 3）。

【0 0 8 3】

例えば、5 分の映像の中から、3 0 秒のシーンを検出したい場合、映像全体からシーンチェンジ検出を行い、その結果を、第 4 のデータ用メモリ 1 1 に記録する。その後、記録されたデータの中から、ちょうど 3 0 秒間隔になっているシーンチェンジの組を探し出し、それを検出結果として出力する。

【0 0 8 4】

図 1 4 を例にとると、シーンチェンジ 1 ～シーンチェンジ 4 が検出された場合に、シーンチェンジ 1 とシーンチェンジ 4 がちょうど 3 0 秒間隔であれば、シー

ンチェンジ1とシーンチェンジ4の組が、検出結果として出力される。

【0085】

このことにより、例えば、テレビ放送などから得られた大量な映像からCM部分のみを取り出したり、放送時間長の決まったニュースや番組を取り出したりすることが可能になる。

【0086】

さらに、あるシーンチェンジが見つかった場合、それから与えられた間隔までのシーンチェンジ判定を省略でき、無駄な検出動作を極力省いて、処理時間を短縮できる。

【0087】

例えば、長い映像の中から、15秒のCMだけを検出したい場合には、ターゲット時間として15秒を与える。シーンチェンジ間隔検索部10では、演算によってシーンチェンジとして検出されたフレームから、15秒後のフレームにシーンチェンジが検出されるかを判定し、15秒後にシーンチェンジが検出された場合にのみ、そのフレームと15秒後のフレームを要求されたシーンチェンジとして出力する。この15秒間がターゲットのシーンとして検出されることになる。続けてターゲットシーンを検索するには、最後に検出されたシーンチェンジからまた15秒後にシーンチェンジが存在するかを判定し、シーンチェンジが検出されなければ次のフレームからシーンチェンジ判定を継続する。検出されればそこがターゲットシーンとなる。

【0088】

図14を例にとると、シーンチェンジ1が検出され、そのちょうど15秒後にシーンチェンジ4があったとすると、この15秒間に存在する、シーンチェンジ2とシーンチェンジ3との判定（無駄な判定）をスキップすることができる。

【0089】

このように、ここでは、一つ目のシーンチェンジが見つかったから、ターゲットの間隔として与えられた時間だけ後のフレームがシーンチェンジであるかどうかを判定している。したがって、検出されたターゲットシーンの中にあるシーンチェンジ判定のための処理を省くことができ、処理時間を短縮できる。



## 【0090】

ここで、本明細書にいう「圧縮動画像のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体」には、複数の記録媒体にプログラムを分散して配布する場合を含む。また、このプログラムが、オペレーティングシステムの一部であるか否かを問わず、種々のプロセスないしスレッド（DLL、OCX、ActiveX等（マイクロソフト社の商標を含む））に機能の一部を肩代わりさせている場合には、肩代わりさせた機能に係る部分が、記録媒体に格納されていない場合も含む。

## 【0091】

図1、図7、図9及び図12（以下「図1等」という）には、スタンドアロン形式のシステムを例示したが、サーバー/クライアント形式にしても良い。つまり、1つの端末機のみ、本明細書に出現する全ての要素が含まれている場合の他、1つの端末機がクライアントであり、これが接続可能なサーバないしネットワーク上に、全部又は一部の要素が実存していても、差し支えない。

## 【0092】

さらには、図1等のほとんどの要素をサーバー側で持ち、クライアント側では、例えば、WWWブラウザだけにしても良い。この場合、各種の情報は、通常サーバ上にあり、基本的にネットワークを経由してクライアントに配布されるものだが、必要な情報が、サーバ上にあるときは、そのサーバの記憶装置が、ここにいる「記録媒体」となり、クライアント上にあるときは、そのクライアントの記録装置が、ここにいる「記録媒体」となる。

## 【0093】

さらに、この「圧縮動画像のシーンチェンジ検出プログラム」には、コンパイルされて機械語になったアプリケーションの他、上述のプロセスないしスレッドにより解釈される中間コードとして実存する場合や、少なくともリソースとソースコードとが「記録媒体」上に格納され、これらから機械語のアプリケーションを生成できるコンパイラ及びリンカが「記録媒体」にある場合や、少なくともリソースとソースコードとが「記録媒体」上に格納され、これらから中間コードのアプリケーションを生成できるインタープリタが「記録媒体」にある場合なども

含む。

【0 0 9 4】

【発明の効果】

請求項 1、2、6、7、11、12 の構成によれば、フレーム構造画像とフィールド構造画像が時間軸上で混在していても、一律にシーンチェンジを検出できる。また、フレーム構造画像内でフレーム D C T を用いられようとフィールド D C T を用いられようと、特別な配慮なしに、シーンチェンジを検出できる。

【0 0 9 5】

請求項 3、8、13 の構成によれば、フィールド符号化ブロック数をカウントし、このカウント数と閾値とを比較することによって、従来検出が非常に困難であった、フィールド間に存在するシーンチェンジをも検出できる。

【0 0 9 6】

請求項 4、9、14 の構成によれば、長い入力動画像から一部のターゲットを容易に取り出し得るし、無駄なシーンチェンジ判定を極力省略でき、その結果、処理時間を短縮できる。

【0 0 9 7】

請求項 5、10、15 の構成によれば、画像の大きさや、データの種類が変わっても、検出のばらつきが少なく、一様の検出結果を得ることができる。

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施の形態におけるシーンチェンジ検出装置のブロック図

【図 2】

本発明の第 1 の実施の形態におけるシーンチェンジ検出装置のフローチャート

【図 3】

同詳細フローチャート

【図 4】

同詳細フローチャート

【図 5】

本発明の第 1 の実施の形態における入力圧縮動画像のモデル図

【図 6】

(a) 本発明の第 1 の実施の形態におけるフィールド構造画像の例示図（トップフィールド）

(b) 同フィールド構造画像の例示図（ボトムフィールド）

(c) 同フレーム構造画像の例示図

【図 7】

本発明の第 2 の実施の形態におけるシーンチェンジ検出装置のブロック図

【図 8】

本発明の第 2 の実施の形態におけるシーンチェンジ検出装置のフローチャート

【図 9】

本発明の第 3 の実施の形態におけるシーンチェンジ検出装置のブロック図

【図 1 0】

本発明の第 3 の実施の形態におけるシーンチェンジ検出装置のフローチャート

【図 1 1】

本発明の第 3 の実施の形態におけるフィールド間に存在するシーンチェンジの例示図

【図 1 2】

本発明の第 4 の実施の形態におけるシーンチェンジ検出装置のブロック図

【図 1 3】

本発明の第 4 の実施の形態におけるシーンチェンジ検出装置のフローチャート

【図 1 4】

本発明の第 4 の実施の形態におけるシーンチェンジのモデル図

【図 1 5】

(a) 従来のフレーム DCT におけるマクロブロック構造の説明図

(b) 従来のフィールド DCT におけるマクロブロック構造の説明図

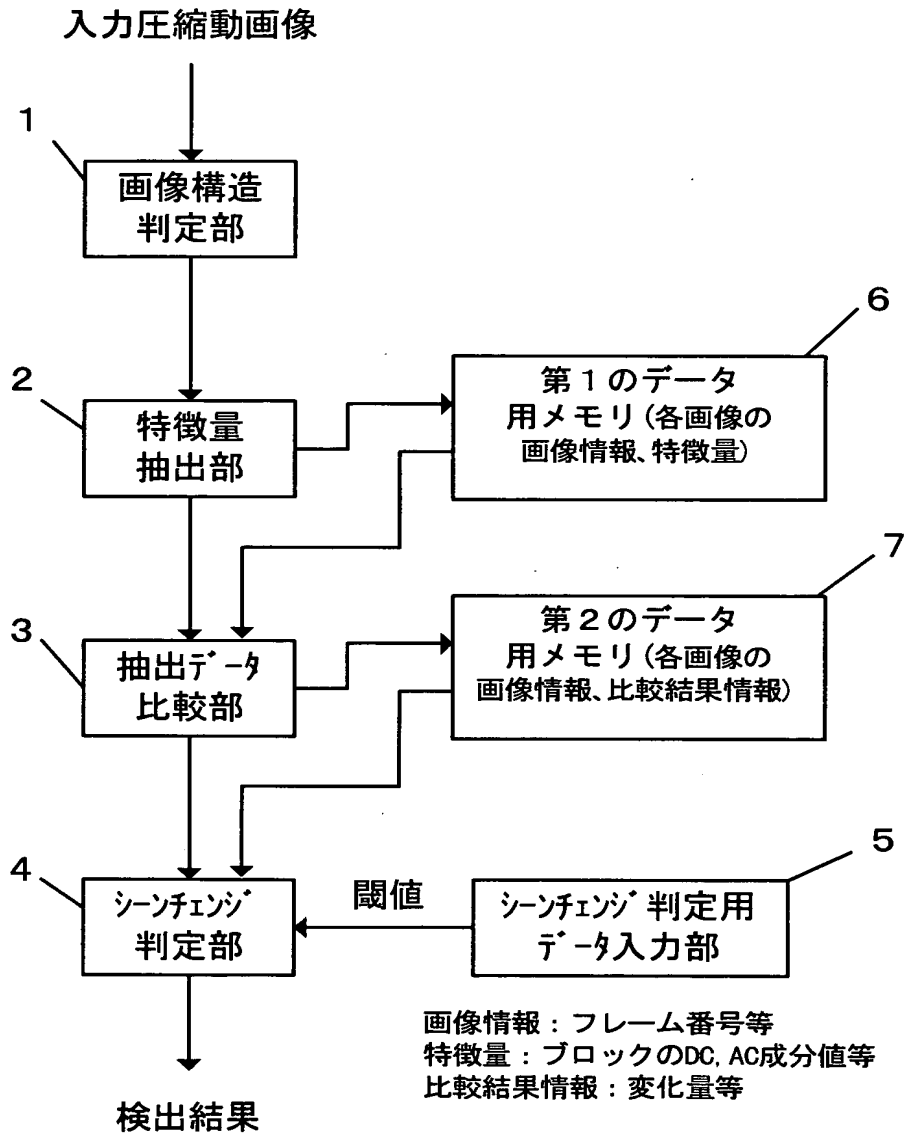
【符号の説明】

- 1 画像構造判定部
- 2 特徴量抽出部
- 3 抽出データ比較部

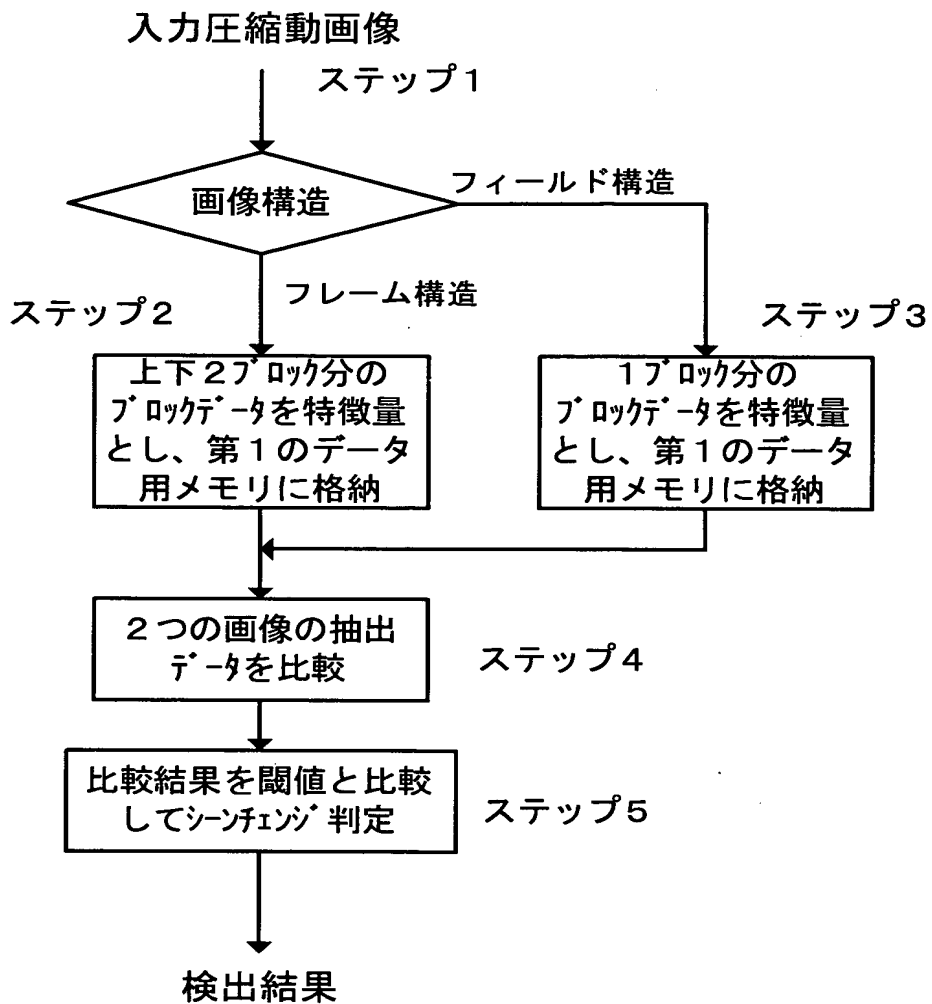
- 4 シーンチェンジ判定部
- 5 シーンチェンジ判定用データ入力部
- 6 第 1 のデータ用メモリ
- 7 第 2 のデータ用メモリ
- 8 フィールド D C T 符号化ブロック数カウント部
- 9 第 3 のデータ用メモリ
- 1 0 シーンチェンジ間隔検索部
- 1 1 第 4 のデータ用メモリ

【書類名】 図面

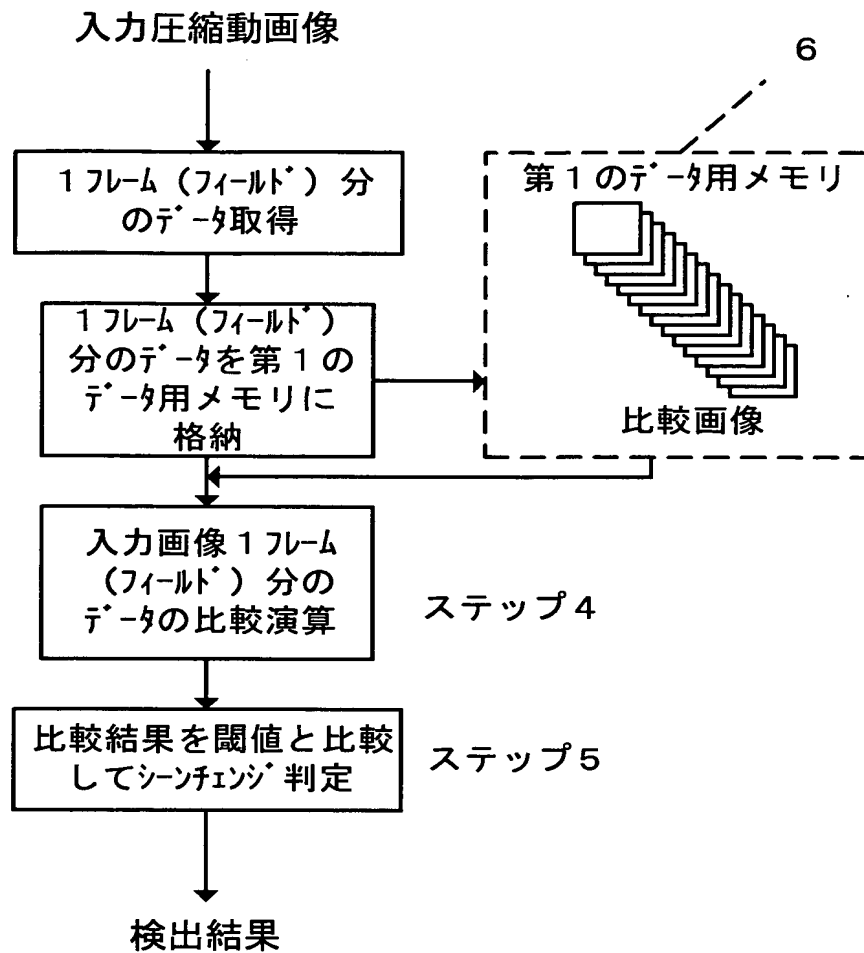
【図 1】



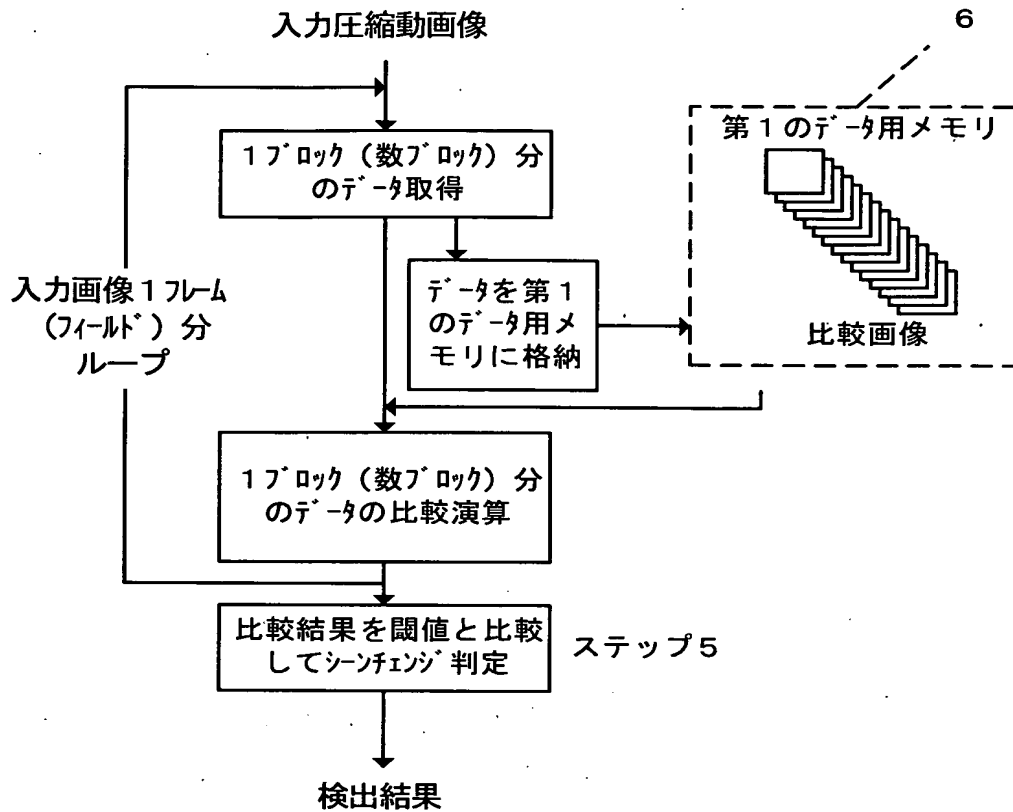
【図 2】



【図 3】

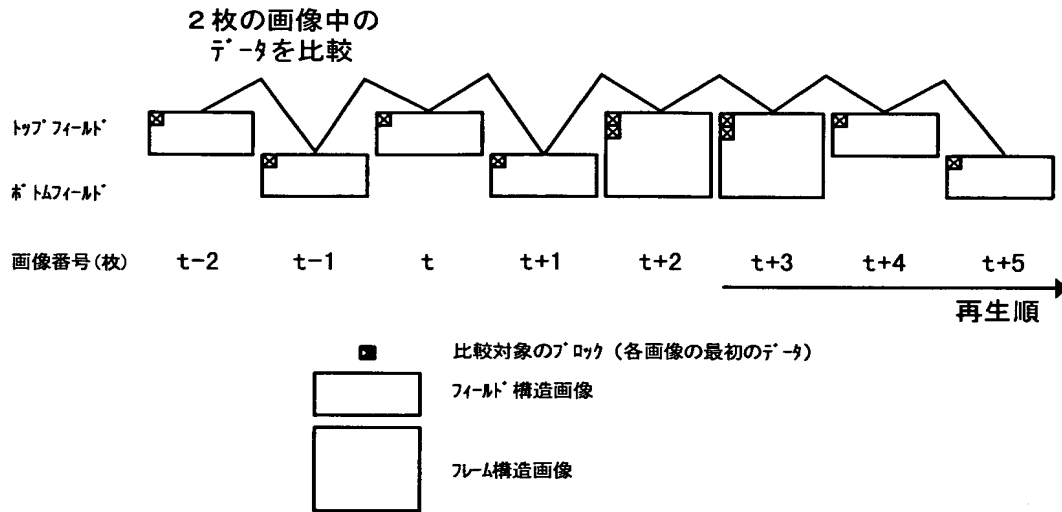


【図 4】

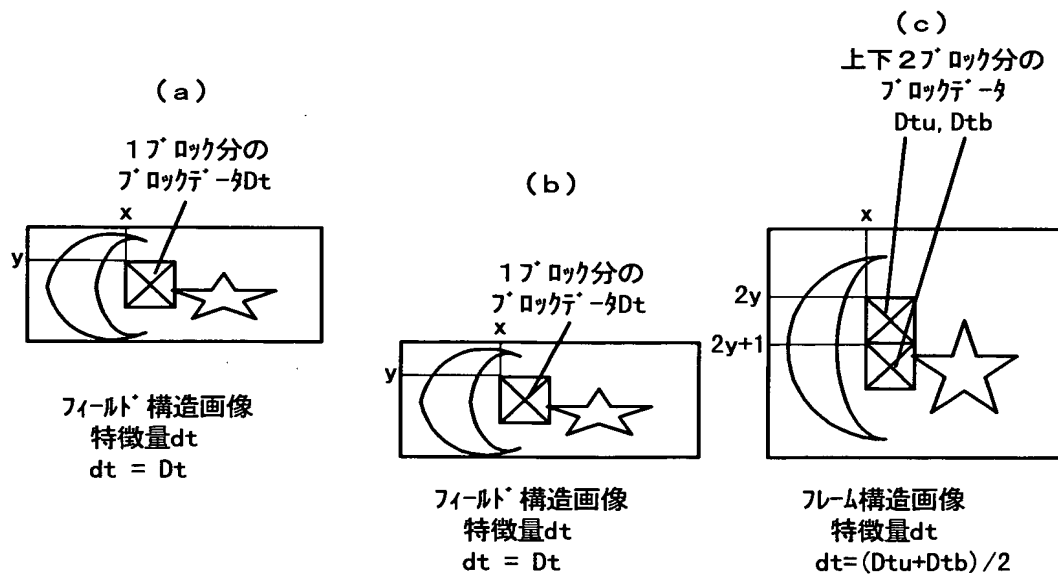




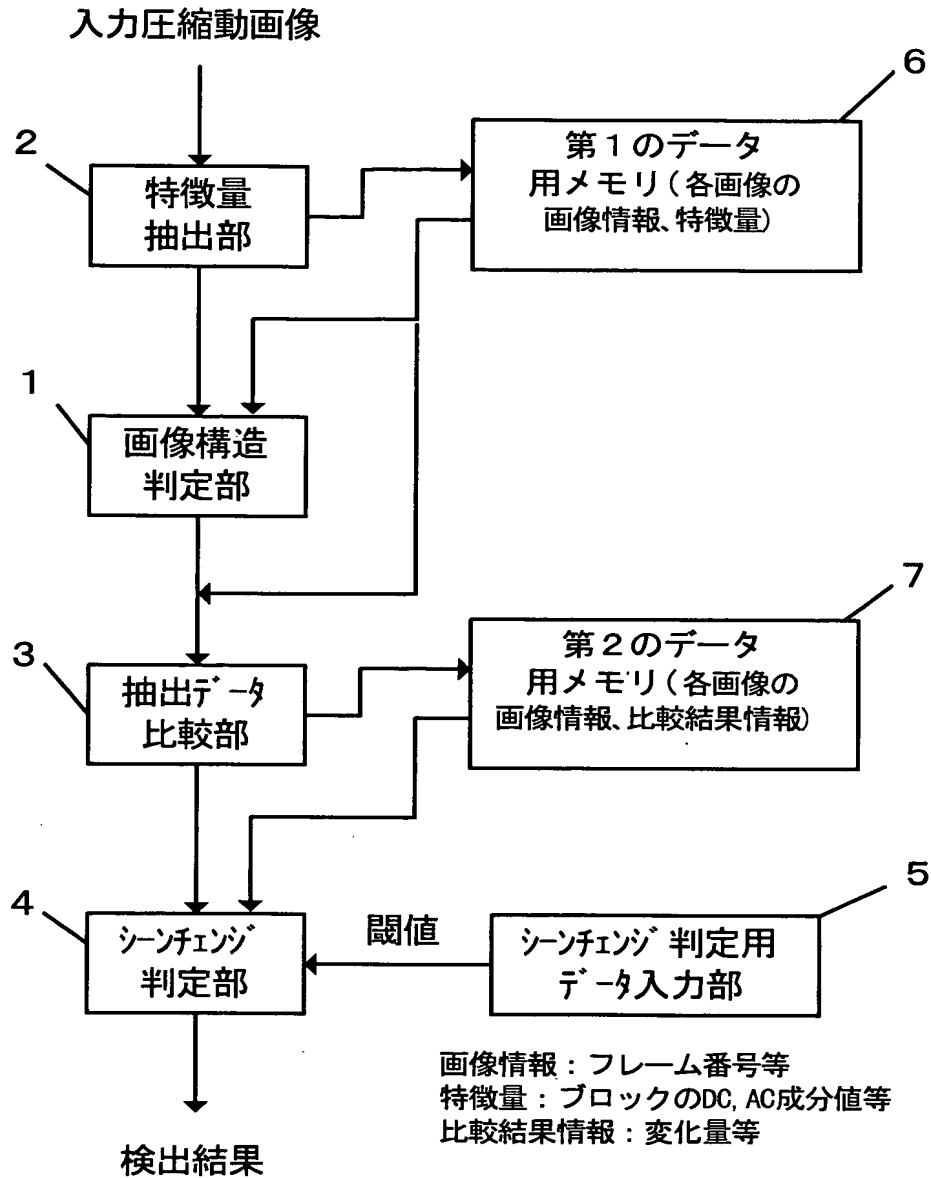
【図 5】



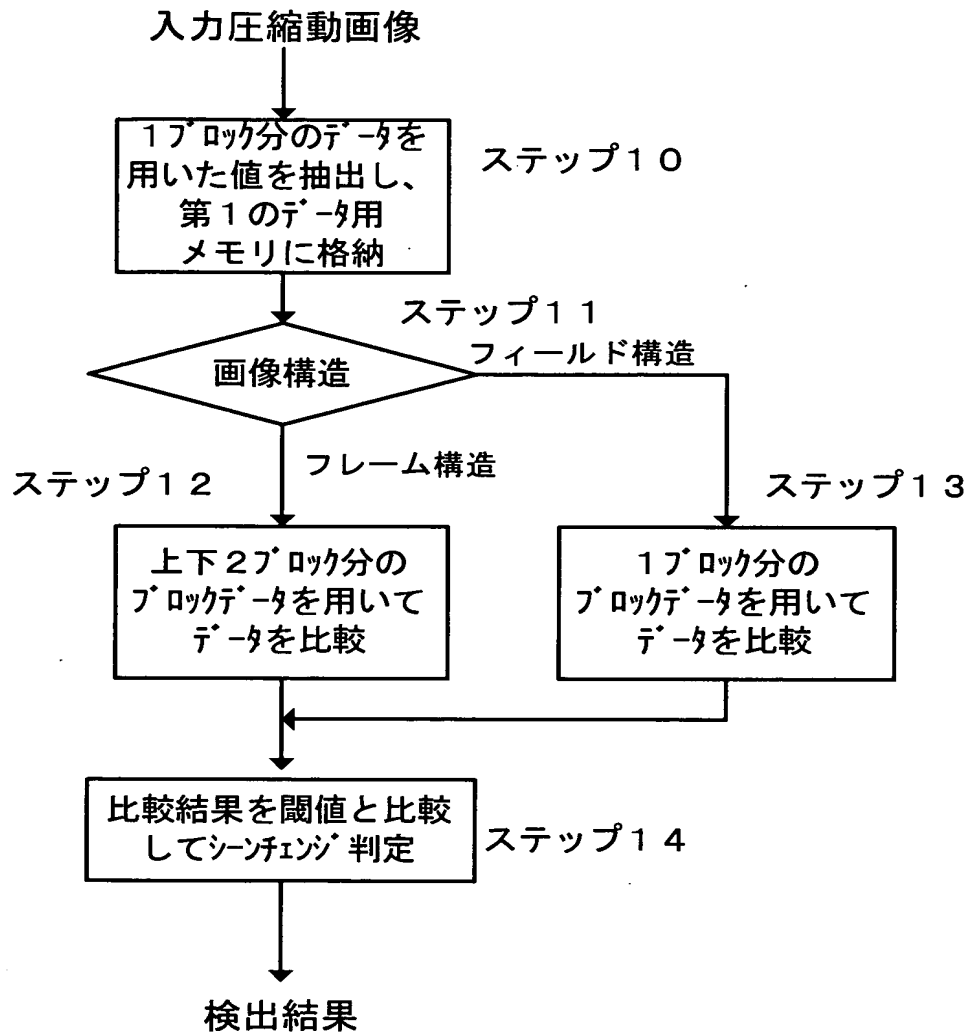
【図 6】



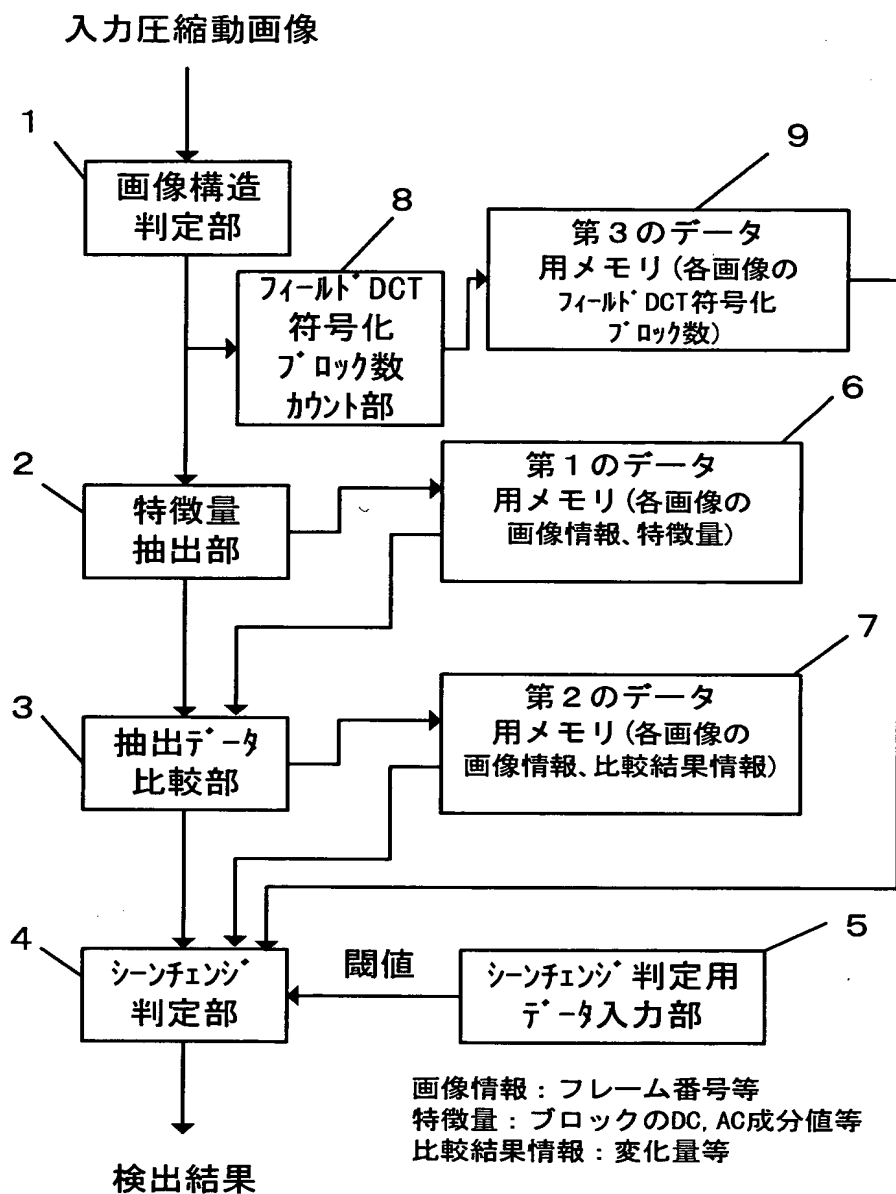
【図 7】



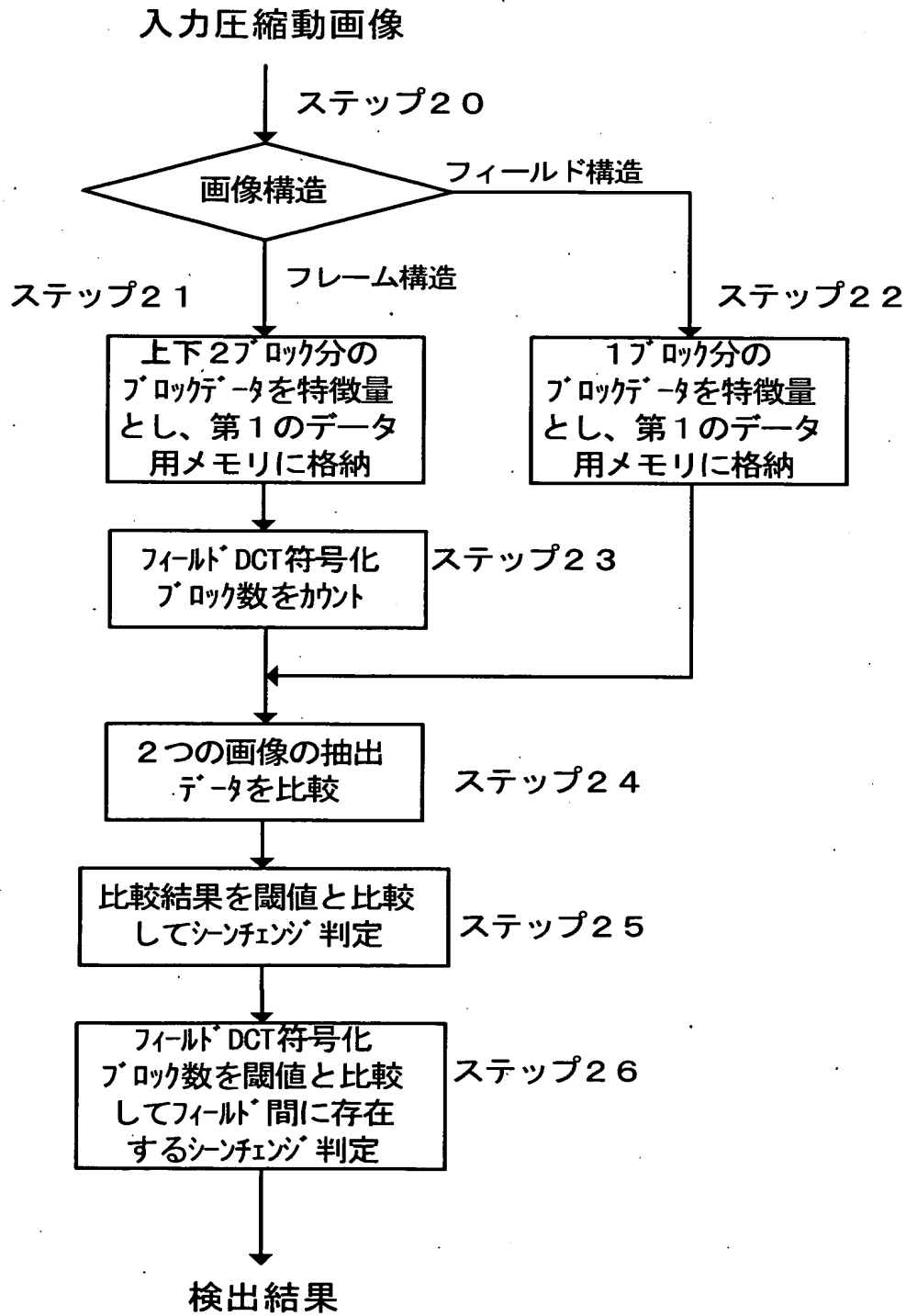
【図 8】



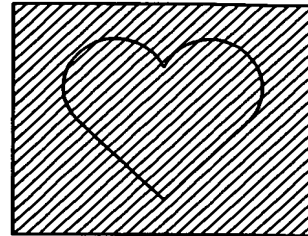
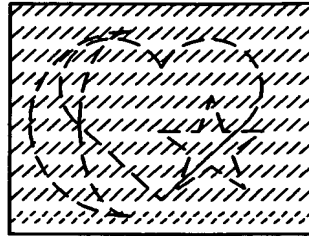
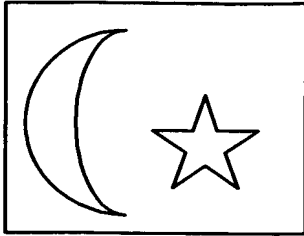
【図 9】



【図 1 0】

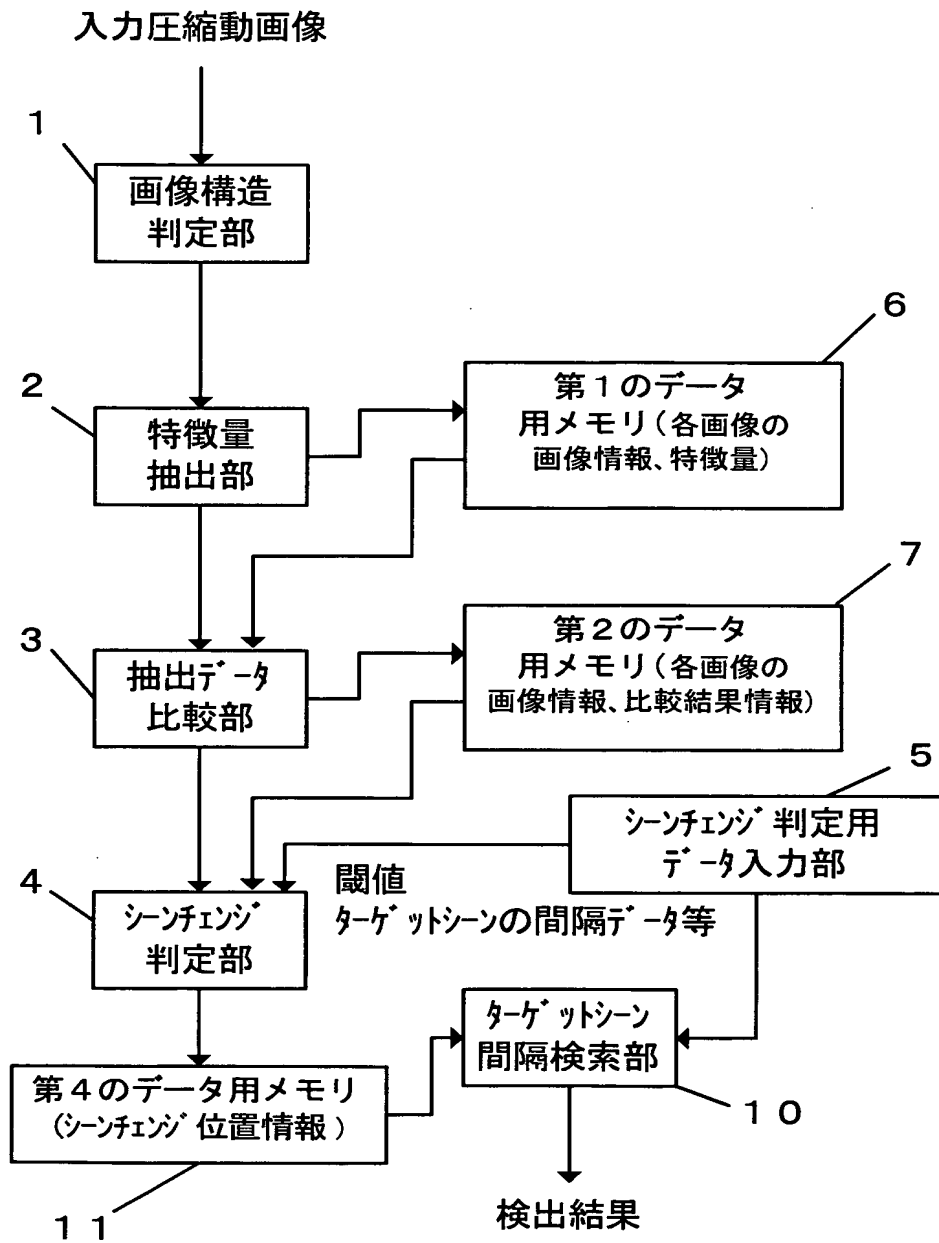


【図 1 1】

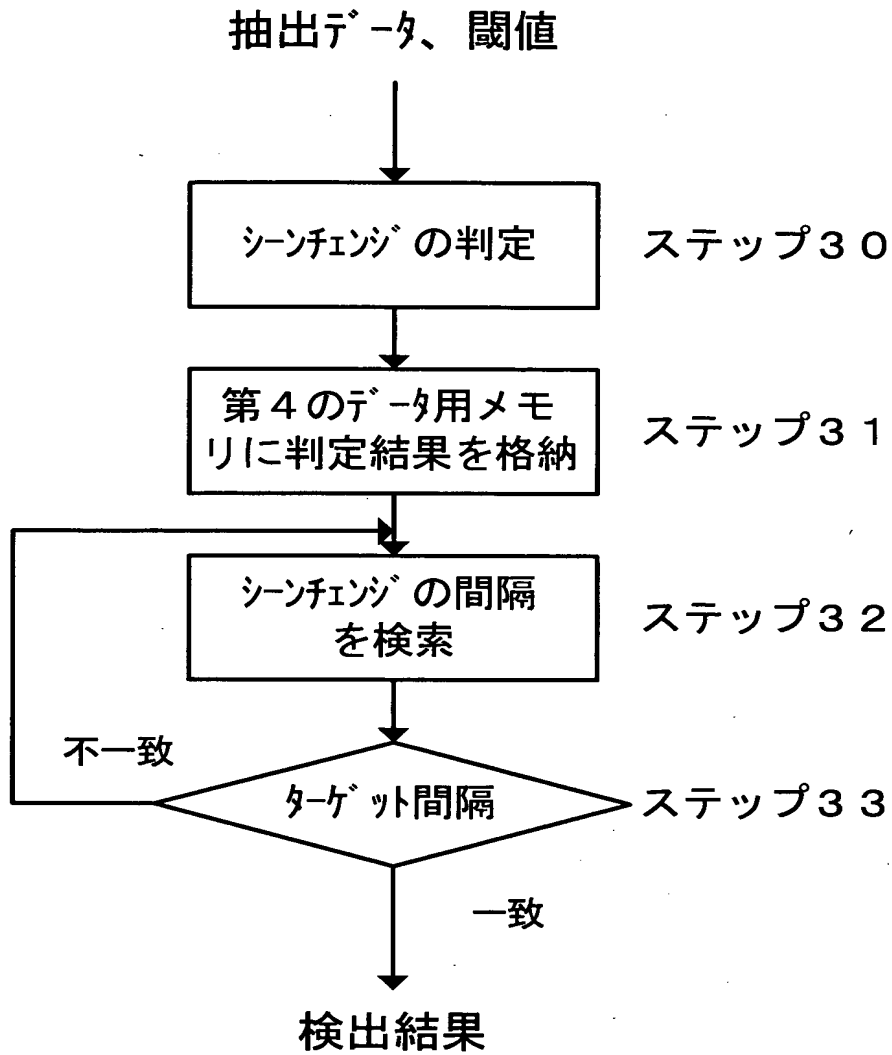


フィールド間での  
シーンチェンジ

【図 12】

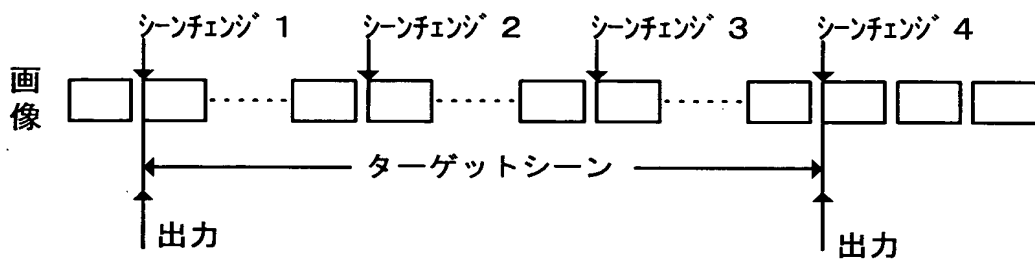


【図 1 3】

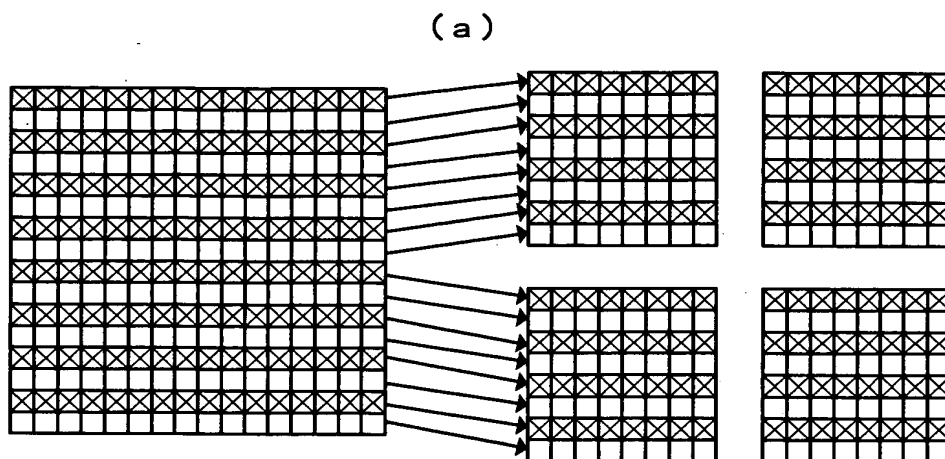




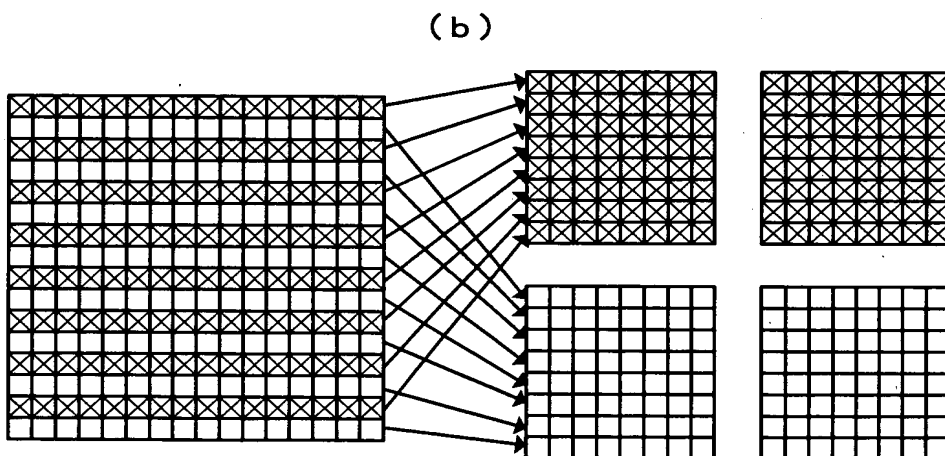
【図 1 4】



【図 1 5】



フレームDCT符号化におけるマクロブロック構造



フィルタDCT符号化におけるマクロブロック構造

【書類名】 要約書

【要約】

【課題】 フレーム／フィールド構造が混在してもシーンチェンジを検出可能な技術を提供する。

【解決手段】 入力圧縮動画像の画像構造を判定する画像構造判定部 1 と、画像構造判定部の判定結果がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に上下 2 倍分のデータに基づいて特徴量を抽出する特徴量抽出部 2 と、特徴量抽出部で抽出されたデータを記録するデータ用メモリ 6 と、抽出されたデータを比較して映像の変化量を求める抽出データ比較部 3 と、抽出データ比較部において求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定部 4 を有する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005821]

1. 変更年月日	1990年 8月28日
[変更理由]	新規登録
住 所	大阪府門真市大字門真1006番地
氏 名	松下電器産業株式会社